1. 第一题

1) 写出贝叶斯最小风险决策和最小错误率决策的决策规则

贝叶斯最小风险决策：

为了最小化总风险，对所有的 i = 1，…，c 计算条件风险：

$$R(\alpha_i \mid \mathbf{x}) = \sum_{j=1}^{c} \lambda_{ij} P(\omega_j \mid \mathbf{x}) = \sum_{j=1}^{c} \lambda_{ij} \frac{P(\mathbf{x} \mid \omega_j) P(\omega_j)}{P(\mathbf{x})}$$

其中α_i是指将分类判别为第 i 类的行为。

则总的条件风险为：

$$\begin{aligned} R &= \int R(\alpha_i \mid \mathbf{x}) P(\mathbf{x}) d\mathbf{x} \\ &= \int \sum_{j=1}^{c} \lambda_{ij} P(\omega_j \mid \mathbf{x}) P(\mathbf{x}) d\mathbf{x} \\ &= \int \sum_{j=1}^{c} \lambda_{ij} \frac{P(\mathbf{x} \mid \omega_j) P(\omega_j)}{P(\mathbf{x})} P(\mathbf{x}) d\mathbf{x} \\ &= \int \sum_{j=1}^{c} \lambda_{ij} P(\mathbf{x} \mid \omega_j) P(\omega_j) d\mathbf{x} \end{aligned}$$

贝叶斯决策的规则就是将上述条件风险最小化

最小错误率决策的决策规则为：

$$\lambda_{ij} = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases} \qquad i, j = 1, \ldots, c$$

$$R(a_i \mid x) = 1 - P(\omega_1 \mid \mathbf{x})$$

则决策规则为：

$$\begin{aligned} R &= \int P(\omega_j \mid \mathbf{x}) P(\mathbf{x}) d\mathbf{x} \\ &= \int \frac{P(\mathbf{x} \mid \omega_j) P(\omega_j)}{P(\mathbf{x})} P(\mathbf{x}) d\mathbf{x} \\ &= \int P(\mathbf{x} \mid \omega_j) P(\omega_j) d\mathbf{x} \end{aligned}$$

2) 写出最小损失决策的决策规则（包括分类规则和拒识规则）

$$\arg\min_i R_i(\mathbf{x}) = \begin{cases} \arg\max_i \dfrac{P(x \mid \omega_i) P(\omega_i)}{P(\mathbf{x})}, & \text{if } \max_i \dfrac{P(x \mid \omega_i) P(\omega_i)}{P(\mathbf{x})} > 1 - \lambda_r / \lambda_s \\ \text{reject} & \text{otherwise} \end{cases}$$

2. 第二题

1) 请写出类条件概率密度函数的数学形式

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp[-\frac{1}{2} (\mathbf{x} - \mu)^t \Sigma^{-1} (\mathbf{x} - \mu)]$$

2) 两种情况下的最小错误率决策判别函数

i. 类协方差矩阵不等

判别函数为：

$$
\begin{aligned}
g_i(\mathbf{x}) &= \mathbf{x}_t \mathbf{W}_i \mathbf{x} + \mathbf{w}_i^t \mathbf{x} + \omega_{i0} \\
\mathbf{W}_i &= -1/2\Sigma_i^{-1} \\
\mathbf{w}_i &= \Sigma_i^{-1}\mu_i \\
\omega_{i0} &= -1/2\mu_i^t\Sigma_i^{-1}\mu_i - 1/2\ln|\Sigma_i| + \ln P(\omega_i)
\end{aligned}
$$

ii. 类协方差矩阵相等

有线性的判别函数：

$$
\begin{aligned}
g_i(\mathbf{x}) &= \mathbf{w}_i^t \mathbf{x} + \omega_{i0} \\
\mathbf{w}_i &= \Sigma^{-1}\mu_i \\
\omega_{i0} &= -\frac{1}{2}\mu_i^t\Sigma^{-1}\mu_i + \ln P(\omega_i)
\end{aligned}
$$

3) 协方差矩阵奇异时的处理方法

i. 协方差矩阵奇异说明必然有两个维度线性相关，因此可以将所有线性相关的维度，只保留一个，即将线性相关的参数去掉至保留一个。

ii. 先假设不同类别的协方差矩阵相同，计算得到共同的协方差矩阵，然后对相同的协方差和不同的协方差之间进行加权平均。

3. 第三题

1) 此规则下的错误率：

由于规则规定一共有两类，不妨分别为w_1和w_2：

则分类为w_1的最小错误率为：

$P(a_1|x) = P(w_2|x)$, 当x小于θ时

则分类为w_2的最小错误率为：

$P(a_2|x) = P(w_1|x)$, 当x大于θ时

则总的错误率为：

$$
\begin{aligned}
& P(error|x) \\
&= P(w_1|x) + P(w_2|x) \\
&= \frac{P(x|w_1)P(w_1) + P(x|w_2)P(w_2)}{P(x)}
\end{aligned}
$$

则$P(error)$

$$
\begin{aligned}
&= P(x|w_1)P(w_1) + P(x|w_2)P(w_2) \\
&= P(w_1)\int_{-\infty}^{\theta} p(x|w_1)dx + P(w_2)\int_{\theta}^{+\infty} p(x|w_2)dx
\end{aligned}
$$

2) 当错误率最小时，满足的条件：

此时 P(error)的导数为 0

$$
P'(error) = P(w_1)p(x|w_1) - P(w_2)p(x|w_2) = 0
$$

$$
P(w_1)p(x|w_1) = P(w_2)p(x|w_2)
$$

3) 这个方程是否决定了θ？

此方程并不能决定θ，因为导数等于 0 的情况除了错误率最小的情况，当错误率最大的时候，导数也等于 0。

4) 举出一个例子

假设样本在(0,1]之间符合均匀分布，且分布在（0, 0.5）之间的类别为 w_2，分布在（0.5, 1]之间的类别为 w_1，在 x=0.5 处 w_1 和 w_2 各占一半，则题目中所述分类规则在θ等于 0.5 的时候取到最大值，而且符合公式：

$$P(w_1)p(x|w_1)=P(w_2)p(x|w_2)$$

4. 第四题

解：

$$p(m,x)=p_m(m)p(x|m)=\frac{1}{2\pi\sigma\sigma_m}\exp(-\frac{\sigma^2(m-m_0)^2+\sigma_m^2(x-m)^2}{2\sigma^2\sigma_m^2})$$

$$p(m|x)=\frac{p(m,x)}{p_x(x)}=\frac{p(m,x)}{\int_R p(m,x)dm}$$

$$\text{有：}\int_{-\infty}^{\infty}\exp(-Am^2+Bm)dm=\int_{-\infty}^{\infty}\exp[-A(m-\frac{B}{2A})^2+\frac{B^2}{4A}]dm$$

$$=\sqrt{\frac{\pi}{A}}\exp(\frac{B^2}{4A})$$

$$\int_R p(m,x)dm=\frac{(2\pi)^{1/2}\sigma\sigma_m}{(\sigma^2+\sigma_m^2)^{\frac{1}{2}}}\exp[\frac{(a_m^2x+m_0\sigma^2)^2}{2\sigma^2\sigma_m^2(\sigma^2+\sigma_m^2)}]$$

$$p(m|x)=\frac{(\sigma^2+\sigma_m^2)^{\frac{1}{2}}}{\sqrt{\pi}\sigma\sigma_m}\exp[-\frac{1}{2}\frac{\sigma^2+\sigma_m^2}{\sigma^2\sigma_m^2}(m-\frac{\sigma_m^2x+m_0\sigma^2}{\sigma^2+\sigma_m^2})^2]$$

5. 编程题：

语言：python 依赖包：numpy

	0	1	2	3	4	5	6	7	8	9	average
LDF	0.963	0.976	0.785	0.881	0.889	0.702	0.914	0.855	0.761	0.797	0.855
QDF	0.935	0.736	0.935	0.894	0.900	0.717	0.890	0.875	0.890	0.822	0.859

总体上来看，QDF 的分类效果略好于 LDF，从各个数字的分类效果上来看，QDF 的大部分数字的分类准确率高于 LDF，但是数字 1 的分类准确率下降比较严重，可能是因为过拟合造成的，测试数据集当中有较多歪斜的数字，导致误判的发生。